

Durham Research Online

Deposited in DRO:

31 August 2016

Version of attached file:

Published Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Cameron, Peter J. and Castillo-Ramirez, Alonso and Gadouleau, Maximilien and Mitchell, James D. (2017) 'Lengths of words in transformation semigroups generated by digraphs.', *Journal of algebraic combinatorics.*, 45 (1). pp. 149-170.

Further information on publisher's website:

<https://doi.org/10.1007/s10801-016-0703-9>

Publisher's copyright statement:

© The Author(s) 2016. Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Additional information:

Use policy


The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

Lengths of words in transformation semigroups generated by digraphs

P. J. Cameron¹ · A. Castillo-Ramirez²  ·
M. Gadouleau² · J. D. Mitchell¹

Received: 2 February 2016 / Accepted: 25 July 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Given a simple digraph D on n vertices (with $n \geq 2$), there is a natural construction of a semigroup of transformations $\langle D \rangle$. For any edge (a, b) of D , let $a \rightarrow b$ be the idempotent of rank $n - 1$ mapping a to b and fixing all vertices other than a ; then, define $\langle D \rangle$ to be the semigroup generated by $a \rightarrow b$ for all $(a, b) \in E(D)$. For $\alpha \in \langle D \rangle$, let $\ell(D, \alpha)$ be the minimal length of a word in $E(D)$ expressing α . It is well known that the semigroup Sing_n of all transformations of rank at most $n - 1$ is generated by its idempotents of rank $n - 1$. When $D = K_n$ is the complete undirected graph, Howie and Iwahori, independently, obtained a formula to calculate $\ell(K_n, \alpha)$, for any $\alpha \in \langle K_n \rangle = \text{Sing}_n$; however, no analogous non-trivial results are known when $D \neq K_n$. In this paper, we characterise all simple digraphs D such that either $\ell(D, \alpha)$ is equal to Howie–Iwahori’s formula for all $\alpha \in \langle D \rangle$, or $\ell(D, \alpha) = n - \text{fix}(\alpha)$ for all $\alpha \in \langle D \rangle$, or $\ell(D, \alpha) = n - \text{rk}(\alpha)$ for all $\alpha \in \langle D \rangle$. We also obtain bounds for $\ell(D, \alpha)$ when D is an acyclic digraph or a strong tournament (the latter case corresponds to a smallest generating set of idempotents of rank $n - 1$ of Sing_n). We finish the paper with a list of conjectures and open problems.

Keywords Transformation semigroup · Simple digraph · Word length

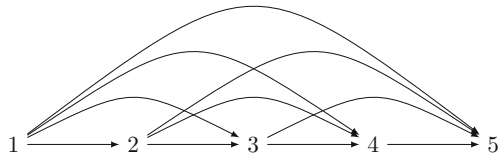
1 Introduction

For any $n \in \mathbb{N}$, $n \geq 2$, let Sing_n be the semigroup of all singular (i.e. non-invertible) transformations on $[n] := \{1, \dots, n\}$. It is well known (see [2]) that Sing_n is generated

✉ A. Castillo-Ramirez
acr8080@gmail.com

¹ School of Mathematics and Statistics, University of St Andrews, St Andrews, Fife KY16 9SS, UK

² School of Engineering and Computing Sciences, Durham University, South Road, Durham DH1 3LE, UK

Fig. 1 \bar{T}_5 

by its idempotents of defect 1 (i.e. the transformations $\alpha \in \text{Sing}_n$ such that $\alpha^2 = \alpha$ and $\text{rk}(\alpha) := |\text{Im}(\alpha)| = n - 1$). There are exactly $n(n - 1)$ such idempotents, and each one of them may be written as $(a \rightarrow b)$, for $a, b \in [n]$, $a \neq b$, where, for any $v \in [n]$,

$$(v)(a \rightarrow b) := \begin{cases} b & \text{if } v = a, \\ v & \text{otherwise.} \end{cases}$$

Motivated by this notation, we refer to these idempotents as *arcs*.

In this paper, we explore the natural connections between simple digraphs on $[n]$ and subsemigroups of Sing_n . For any subset $U \subseteq \text{Sing}_n$, denote by $\langle U \rangle$ the semigroup generated by U . For any simple digraph D with vertex set $V(D) = [n]$ and edge set $E(D)$, we associate the semigroup

$$\langle D \rangle := \langle (a \rightarrow b) \in \text{Sing}_n : (a, b) \in E(D) \rangle.$$

We say that a subsemigroup S of Sing_n is *arc-generated* by a simple digraph D if $S = \langle D \rangle$.

For the rest of the paper, we use the term ‘digraph’ to mean ‘simple digraph’ (i.e. a digraph with no loops or multiple edges). A digraph D is *undirected* if its edge set is a symmetric relation on $V(D)$, and it is *transitive* if its edge set is a transitive relation on $V(D)$. We shall always assume that D is *connected* (i.e. for every pair $u, v \in V(D)$ there is either a path from u to v , or a path from v to u) because otherwise $\langle D \rangle \cong \langle D_1 \rangle \times \cdots \times \langle D_k \rangle$, where D_1, \dots, D_k are the connected components of D . We say that D is *strong* (or *strongly connected*) if for every pair $u, v \in V(D)$, there is a directed path from u to v . We say that D is a *tournament* if for every pair $u, v \in V(D)$ we have $(u, v) \in E(D)$ or $(v, u) \in E(D)$, but not both.

Many famous examples of semigroups are arc-generated. Clearly, by the discussion of the first paragraph, Sing_n is arc-generated by the complete undirected graph K_n . In fact, for $n \geq 3$, Sing_n is arc-generated by D if and only if D contains a strong tournament (see [3]). The semigroup of order-preserving transformations $\text{O}_n := \{\alpha \in \text{Sing}_n : u \leq v \Rightarrow u\alpha \leq v\alpha\}$ is arc-generated by an undirected path P_n on $[n]$, while the Catalan semigroup $\text{C}_n := \{\alpha \in \text{Sing}_n : v \leq v\alpha, u \leq v \Rightarrow u\alpha \leq v\alpha\}$ is arc-generated by a directed path \bar{P}_n on $[n]$ (see [9, Corollary 4.11]). The semigroup of non-decreasing transformations $\text{OI}_n := \{\alpha \in \text{Sing}_n : v \leq v\alpha\}$ is arc-generated by the transitive tournament \bar{T}_n on $[n]$ (Fig. 1 illustrates \bar{T}_5).

Connections between subsemigroups of Sing_n and digraphs have been studied before (see [9–12]). The following definition, which we shall adopt in the following sections, appeared in [12]:

Definition 1 For a digraph D , the *closure* \bar{D} of D is the digraph with vertex set $V(\bar{D}) := V(D)$ and edge set $E(\bar{D}) := E(D) \cup \{(a, b) : (b, a) \in E(D) \text{ is in a directed cycle of } D\}$.

Say that D is *closed* if $D = \bar{D}$. Observe that $\langle D \rangle = \langle \bar{D} \rangle$ for any digraph D .

Recall that the *orbits* of $\alpha \in \text{Sing}_n$ are the connected components of the digraph on $[n]$ with edges $\{(x, x\alpha) : x \in [n]\}$. In particular, an orbit Ω of α is called *cyclic* if it is a cycle with at least two vertices. An element $x \in [n]$ is a *fixed point* of α if $x\alpha = x$. Denote by $\text{cycl}(\alpha)$ and $\text{fix}(\alpha)$ the number of cyclic orbits and fixed points of α , respectively. Denote by $\ker(\alpha)$ the partition of $[n]$ induced by the *kernel* of α (i.e. the equivalence relation $\{(x, y) \in [n]^2 : x\alpha = y\alpha\}$).

We introduce some further notation. For any digraph D and $v \in V(D)$, define the *in-neighbourhood* and the *out-neighbourhood* of v by

$$N^-(v) := \{u \in V(D) : (u, v) \in E(D)\} \text{ and } N^+(v) := \{u \in V(D) : (v, u) \in E(D)\},$$

respectively. We extend these definitions to any subset $C \subseteq V(D)$ by letting $N^\epsilon(C) := \bigcup_{c \in C} N^\epsilon(c)$, where $\epsilon \in \{+, -\}$. The *in-degree* and *out-degree* of v are $\deg^-(v) := |N^-(v)|$ and $\deg^+(v) := |N^+(v)|$, respectively, while the *degree* of v is $\deg(v) := |N^-(v) \cup N^+(v)|$. For any two vertices $u, v \in V(D)$, the D -*distance* from u to v , denoted by $d_D(u, v)$, is the length of a shortest path from u to v in D , provided that such a path exists. The *diameter* of D is $\text{diam}(D) := \max\{d_D(u, v) : u, v \in V(D), d_D(u, v) \text{ is defined}\}$.

Let D be any digraph on $[n]$. We are interested in the lengths of transformations of $\langle D \rangle$ viewed as words in the free monoid $D^* := \{(a \rightarrow b) : (a, b) \in E(D)\}^*$. Say that a word $\omega \in D^*$ *expresses* (or *evaluates to*) $\alpha \in \langle D \rangle$ if $\alpha = \omega\phi$, where $\phi : D^* \rightarrow \langle D \rangle$ is the evaluation semigroup morphism. For any $\alpha \in \langle D \rangle$, let $\ell(D, \alpha)$ be the minimum length of a word in D^* expressing α . For $r \in [n - 1]$, denote

$$\begin{aligned} \ell(D, r) &:= \max \{\ell(D, \alpha) : \alpha \in \langle D \rangle, \text{rk}(\alpha) = r\}, \\ \ell(D) &:= \max \{\ell(D, \alpha) : \alpha \in \langle D \rangle\}. \end{aligned}$$

The main result in the literature in the study of $\ell(D, \alpha)$ was obtained by Howie and Iwahori, independently, when $D = K_n$.

Theorem 1.1 [4,5] *For any $\alpha \in \text{Sing}_n$,*

$$\ell(K_n, \alpha) = n + \text{cycl}(\alpha) - \text{fix}(\alpha).$$

Therefore, $\ell(K_n, r) = n + \lfloor \frac{1}{2}(r - 2) \rfloor$, for any $r \in [n - 1]$, and $\ell(K_n) = \ell(K_n, n - 1) = \lfloor \frac{3}{2}(n - 1) \rfloor$.

In the following sections, we study $\ell(D, \alpha)$, $\ell(D, r)$, and $\ell(D)$, for various classes of digraphs. In Sect. 2, we characterise all digraphs D on $[n]$ such that either $\ell(D, \alpha) = n + \text{cycl}(\alpha) - \text{fix}(\alpha)$ for all $\alpha \in \langle D \rangle$, or $\ell(D, \alpha) = n - \text{fix}(\alpha)$ for all $\alpha \in \langle D \rangle$, or $\ell(D, \alpha) = n - \text{rk}(\alpha)$ for all $\alpha \in \langle D \rangle$. In Sect. 3, we are interested in the maximal possible length of a transformation in $\langle D \rangle$ of rank r among all digraphs D on $[n]$ of

certain class \mathcal{C} ; we denote this number by $\ell_{\max}^{\mathcal{C}}(n, r)$. In particular, when \mathcal{C} is the class of acyclic digraphs, we find an explicit formula for $\ell_{\max}^{\mathcal{C}}(n, r)$. When \mathcal{C} is the class of strong tournaments, we find upper and lower bounds for $\ell_{\max}^{\mathcal{C}}(n, r)$ (and for the analogously defined $\ell_{\min}^{\mathcal{C}}(n, r)$). Finally, in Sect. 4 we provide a list of conjectures and open problems.

2 Arc-generated semigroups with short words

Let D be a digraph on $[n]$, $n \geq 3$, and $\alpha \in \langle D \rangle$. Theorem 1.1 implies the following three bounds:

$$\ell(D, \alpha) \geq n + \text{cycl}(\alpha) - \text{fix}(\alpha) \geq n - \text{fix}(\alpha) \geq n - \text{rk}(\alpha). \quad (1)$$

The lowest bound is always achieved for constant transformations (i.e. transformations of rank 1).

Lemma 2.1 *For any digraph D on $[n]$, if $\alpha \in \langle D \rangle$ has rank 1, then $\ell(D, \alpha) = n - 1$.*

Proof It is clear that $\ell(D, \alpha) \geq n - 1$ because α has $n - 1$ non-fixed points. Let $\text{Im}(\alpha) = \{v_0\} \subseteq [n]$. Note that, for any $v \in [n]$, there is a directed path in D from v to v_0 (as otherwise, $\alpha \notin \langle D \rangle$). For any $d \geq 1$, let

$$C_d := \{v \in [n] : d_D(v, v_0) = d\}.$$

Clearly, $[n] \setminus \{v_0\} = \bigcup_{d=1}^m C_d$, where $m := \max_{v \in [n]} \{d_D(v, v_0)\}$ and the union is disjoint. For any $v \in C_d$, let v' be a vertex in C_{d-1} such that $(v \rightarrow v') \in D$. For any distinct $v, u \in C_d$ and any choice of $v', u' \in C_{d-1}$, the arcs $(v \rightarrow v')$ and $(u \rightarrow u')$ commute; hence, we can decompose α as

$$\alpha = \bigcirc_{d=m}^1 \bigcirc_{v \in C_d} (v \rightarrow v'),$$

where the composition of arcs is done from m down to 1. \square

Remark 1 Using a similar argument as in the previous proof, we may show that $\langle D \rangle$ contains all constant transformations if and only if D is strongly connected.

Inspired by the bounds given in (1), we characterise all the connected digraphs D on $[n]$ satisfying the following conditions:

$$\forall \alpha \in \langle D \rangle, \ell(D, \alpha) = n + \text{cycl}(\alpha) - \text{fix}(\alpha); \quad (\text{C1})$$

$$\forall \alpha \in \langle D \rangle, \ell(D, \alpha) = n - \text{fix}(\alpha); \quad (\text{C2})$$

$$\forall \alpha \in \langle D \rangle, \ell(D, \alpha) = n - \text{rk}(\alpha). \quad (\text{C3})$$

2.1 Digraphs satisfying condition (C1)

Theorem 1.1 says that K_n satisfies (C1). In order to characterise all digraphs satisfying (C1), we introduce the following property on a digraph D :

- (\star) If $d_D(v_0, v_2) = 2$ and v_0, v_1, v_2 is a directed path in D , then $N^+(\{v_1, v_2\}) \subseteq \{v_0, v_1, v_2\}$.

We shall study the strong components of digraphs satisfying property (\star). We state few observations that we use repeatedly in this section.

Remark 2 Suppose that D satisfies property (\star). If v_0, v_1, v_2 is a directed path in D and $\deg^+(v_1) > 2$, or $\deg^+(v_2) > 2$, then $(v_0, v_2) \in E(D)$. Indeed, if $(v_0, v_2) \notin E(D)$, then $d_D(v_0, v_2) = 2$, so, by property (\star), $N^+(\{v_1, v_2\}) \subseteq \{v_0, v_1, v_2\}$; this contradicts that $\deg^+(v_1) > 2$, or $\deg^+(v_2) > 2$.

Remark 3 Suppose that D satisfies property (\star). If v_0, v_1, v_2 is a directed path in D and either v_1 or v_2 has an out-neighbour not in $\{v_0, v_1, v_2\}$, then $(v_0, v_1) \in E(D)$.

Remark 4 If D satisfies property (\star), then $\text{diam}(D) \leq 2$. Indeed, if v_0, v_1, \dots, v_k is a directed path in D with $d_D(v_0, v_k) = k \geq 3$, then v_0, v_1, v_2 is a directed path in D and v_2 has an out-neighbour $v_3 \notin \{v_0, v_1, v_2\}$; by Remark 3, $(v_0, v_2) \in E(D)$, which contradicts that $d_D(v_0, v_k) = k$.

Note that digraphs satisfying property (\star) are a slight generalisation of transitive digraphs.

Let D be a digraph and let C_1 and C_2 of be two strong components of D . We say that C_1 *connects* to C_2 if $(v_1, v_2) \in E(D)$ for some $v_1 \in C_1, v_2 \in C_2$; similarly, we say that C_1 *fully connects* to C_2 if $(v_1, v_2) \in E(D)$ for all $v_1 \in C_1, v_2 \in C_2$. The strong component C_1 is called *terminal* if there is no strong component $C \neq C_1$ of D such that C_1 connects to C .

Lemma 2.2 *Let D be a closed digraph satisfying property (\star). Then, any strong component of D is either an undirected path P_3 or complete. Furthermore, P_3 may only appear as a terminal strong component of D .*

Proof Let C be a strong component of D . Since D is closed, C must be undirected. The lemma is clear if $|C| \leq 3$, so assume that $|C| \geq 4$. We have two cases:

Case 1 Every vertex in C has degree at most 2. Then C is a path or a cycle. Since $|C| \geq 4$ and $\text{diam}(D) \leq 2$, then C is a cycle of length 4 or 5; however, these cycles do not satisfy property (\star).

Case 2 There exists a vertex $a \in C$ of degree 3 or more. Any two neighbours of a are adjacent: indeed, for any $u, v \in N(a)$, u, a, v is a path and $\deg^+(a) > 2$, so $(u, v) \in E(D)$ by Remark 2. Hence, the neighbourhood of a is complete and every neighbour of a has degree 3 or more. Applying this rule recursively, we obtain that every vertex in C has degree 3 or more, and the neighbourhood of every vertex is complete. Therefore, C is complete because $\text{diam}(D) \leq 2$.

Finally, if P_3 is a strong component of D , there cannot be any edge coming out of it because of property (\star), so it must be a terminal component. \square

Lemma 2.3 *Let D be a closed digraph satisfying property (\star) . Let C_1 and C_2 be strong components of D , and suppose that C_1 connects to C_2 .*

- (i) *If C_2 is non-terminal, then C_1 fully connects to C_2 .*
- (ii) *Let $|C_2| = 1$. If either $|C_1| \neq 2$, or the vertex in C_1 that connects to C_2 has out-degree at least 3, then C_1 fully connects to C_2 .*
- (iii) *Let $|C_2| = 2$. If not all vertices in C_1 connect to the same vertex in C_2 , then C_1 fully connects to C_2 .*
- (iv) *If $|C_2| \geq 3$, then C_1 fully connects to C_2 .*

Proof Recall that C_1 and C_2 are undirected because D is closed. If $|C_1| = 1$ and $|C_2| = 1$, clearly C_1 fully connects to C_2 . Henceforth, we assume $|C_1| \geq 2$ or $|C_2| \geq 2$. Let $c_1 \in C_1$ and $c_2 \in C_2$ be such that $(c_1, c_2) \in E(D)$. As C_1 is a non-terminal, Lemma 2.2 implies that C_1 is complete.

- (i) As C_2 is non-terminal, there exists $d \in D \setminus (C_1 \cup C_2)$ such that $(c_2, d) \in E(D)$. Suppose that $|C_1| \geq 2$. Then, for any $c'_1 \in C_1 \setminus \{c_1\}$, c'_1, c_1, c_2 is a directed path in D with $d \in N^+(c_2)$, so Remark 3 implies $(c'_1, c_2) \in E(D)$. Suppose now that $|C_2| \geq 2$. Then, for any $c'_2 \in C_2 \setminus \{c_2\}$, c_1, c_2, c'_2 is a directed path in D with $d \in N^+(c_2)$, so again $(c_1, c'_2) \in E(D)$. Therefore, C_1 fully connects to C_2 .
- (ii) Suppose that $|C_1| \geq 2$. If $|C_1| > 2$, then $\deg^+(c_1) > 2$, because C_1 is complete. Thus, for each $c'_1 \in C_1 \setminus \{c_1\}$, c'_1, c_1, c_2 is a directed path in D with $\deg^+(c_1) > 2$, so $(c'_1, c_2) \in E(D)$ by Remark 2. As $|C_2| = 1$, this shows that C_1 fully connects to C_2 .
- (iii) Let $C_2 = \{c_2, c'_2\}$ and let $c'_1 \in C_1 \setminus \{c_1\}$ be such that $(c'_1, c'_2) \in E(D)$. For any $b, d \in C_1$, $b \neq c_1$, $d \neq c'_1$, both b, c_1, c_2 and d, c'_1, c'_2 are directed paths in D with $c'_2 \in N^+(c_2)$ and $c_2 \in N^+(c'_2)$; hence, $(b, c_2), (d, c'_2) \in E(D)$ by Remark 3.
- (iv) Suppose that $C_2 = P_3$. Say $C_2 = \{c_2, c'_2, c''_2\}$ with either $d_D(c_2, c'_2) = 2$ or $d_D(c'_2, c''_2) = 2$. In any case, c_1, c_2, c'_2 is a directed path in D with $c''_2 \in N^+(\{c_2, c'_2\})$, so $(c_1, c'_2) \in E(D)$ by Remark 3; now, c_1, c'_2, c''_2 is a directed path in D with $c_2 \in N^+(\{c'_2, c''_2\})$, so $(c_1, c''_2) \in E(D)$. Hence, c_1 is connected to all vertices of C_2 . As C_1 is complete, a similar argument shows that every $c'_1 \in C_1 \setminus \{c_1\}$ connects to every vertex in C_2 .
Suppose now that $C_2 = K_m$ for $m \geq 3$. By a similar reasoning as the previous paragraph, we show that $(c_1, v) \in E(D)$ for all $v \in C_2$. Now, for any $c'_1 \in C_1 \setminus \{c_1\}$, $v \in C_2$, c'_1, c_1, v is a directed path in D so $(c'_1, v) \in E(D)$ by Remark 3. \square

Lemma 2.4 *Let D be a closed digraph satisfying property (\star) . Let C_i , $i = 1, 2, 3$, be strong components of D , and suppose that C_1 connects to C_2 and C_2 connects to C_3 . If C_1 does not connect to C_3 , then $|C_2| = |C_3| = 1$, C_3 is terminal in D , and C_2 is terminal in $D \setminus C_3$.*

Proof By Lemma 2.3 (i), C_1 fully connects to C_2 . Assume that C_1 does not connect to C_3 . Let $c_i \in C_i$, $i = 1, 2, 3$, be such that $(c_1, c_2), (c_2, c_3) \in E(D)$. If C_2 has a vertex different from c_2 , Remark 3 ensures that $(c_1, c_3) \in E(D)$, which contradicts our hypothesis. Then $|C_2| = 1$. The same argument applies if C_3 has a vertex different

from c_3 , so $|C_3| = 1$. Finally, Remark 3 applied to the path c_1, c_2, c_3 also implies that C_3 is terminal in D and C_2 is terminal in $D \setminus C_3$. \square

The following result characterises all digraphs satisfying condition (C1).

Theorem 2.5 *Let D be a connected digraph on $[n]$. The following are equivalent:*

- (i) *For all $\alpha \in \langle D \rangle$, $\ell(D, \alpha) = n + \text{cycl}(\alpha) - \text{fix}(\alpha)$.*
- (ii) *D is closed satisfying property (\star) .*

Proof In order to simplify notation, denote

$$g(\alpha) := n + \text{cycl}(\alpha) - \text{fix}(\alpha).$$

First, we show that (i) implies (ii). Suppose $\ell(D, \alpha) = g(\alpha)$ for all $\alpha \in \langle D \rangle$. We use the one-line notation for transformations: $\alpha = (1)\alpha (2)\alpha \dots (k)\alpha$, where $x = (x)\alpha$ for all $x > k$, $x \in [n]$. Clearly, if D is not closed, there exists an arc $\alpha \in \langle D \rangle \setminus D$, so $1 < \ell(D, \alpha) \neq g(\alpha) = 1$. In order to prove that property (\star) holds, let 1, 2, 3 be a shortest path in D . If $(2 \rightarrow v) \in \langle D \rangle$, for some $v \in [n] \setminus \{1, 2, 3\}$, then $\alpha = 3v3v \in \langle D \rangle$, but $g(\alpha) = 2 \neq \ell(D, \alpha) = 3$. If $(3 \rightarrow v) \in \langle D \rangle$, then $\alpha = 3vvv \in \langle D \rangle$, but $g(\alpha) = 3 \neq \ell(D, \alpha) = 4$. Therefore, $N^+(\{2, 3\}) \subseteq \{1, 2, 3\}$, and (\star) holds.

Conversely, we show that (ii) implies (i). Let $\alpha \in \langle D \rangle$. We remark that any cycle of α belongs to a strong component of D .

Claim 2.6 *Let C be a strong component of D . Then either α fixes all vertices of C or $|(C\alpha) \cap C| < |C|$.*

Proof Suppose that $\alpha|_C$, the restriction of α to C , is non-trivial and $|(C\alpha) \cap C| = |C|$. Then $\alpha|_C$ is a permutation of C . Let $u \in C$ and suppose that $(u \rightarrow v)$ is the first arc moving u in a word expressing α in D^* . If $v \in C$, we have $u\alpha = v\alpha$, which contradicts that $\alpha|_C$ is a permutation. If $v \in C'$ for some other strong component C' of D , then $u\alpha \notin C$ which again contradicts our assumption. \square

Claim 2.7 *Let $u, v \in [n]$ be such that $u\alpha = v$. If $d_D(u, v) = 2$, then:*

- 1. *v is in a terminal component of D .*
- 2. *There is a path u, w, v of length 2 in D such that $w\alpha = v\alpha = v$; for any other path u, x, v of length 2 in D , we have $x\alpha \in \{x, v\}$.*

Proof Let C_1 and C_2 be strong components of D such that $u \in C_1$ and $v \in C_2$. We analyse the four possible cases in which $d_D(u, v) = 2$. In the first three cases, we use the fact that $\langle P_3 \rangle \cong O_3$, hence we can order $u < w < v$ and α is an increasing transformation of the ordered set $\{u, w, v\}$; thus $u\alpha = w\alpha = v\alpha = v$.

Case 1 $C_1 = C_2$. By Lemma 2.2, $C_1 \cong P_3$ and it is a terminal component. Therefore, 2. holds as there is a unique path from u to v .

Case 2 C_1 connects to C_2 and $|C_2| \neq 2$. As $d_D(u, v) = 2$, C_1 does not fully connect C_2 , so, by Lemma 2.3, $|C_2| = 1$, C_2 is terminal, $|C_1| = 2$, and the vertex $w \in C_1$ connecting to $C_2 = \{v\}$ has out-degree 2. Then, by property (\star) , u, w, v is the unique path from u to v .

- Case 3** C_1 connects to C_2 and $|C_2| = 2$. As $d_D(u, v) = 2$, C_1 does not fully connect C_2 , so, by Lemma 2.3, C_2 is terminal and u, w, v is the unique path of length two from u to v , where w is the other vertex of C_2 .
- Case 4** C_1 does not connect to C_2 . Since $d_D(u, v) = 2$, there exist strong components $C^{(1)}, \dots, C^{(k)}$ such that C_1 connects to $C^{(i)}$ and $C^{(i)}$ connects to C_2 , for all $1 \leq i \leq k$. By Lemma 2.4, $C^{(i)} = \{x_i\}$, $C_2 = \{v\}$ is terminal and $N^+(x_i) = \{v\}$ for all i . Thus u, x_i, v are the only paths of length two from u to v ; in particular, $x_i\alpha \in \{x_i, v\}$ for all x_i . As $u\alpha = v$, there must exist $1 \leq j \leq k$ such that $w := x_j$ is mapped to v . \square

Now we produce a word $\omega \in D^*$ expressing α of length $g(\alpha)$. Define

$$U := \{u \in D : d_D(u, u\alpha) = 2\}.$$

For every $u \in U$, let u' be a vertex in D such that $u, u', u\alpha$ is a path and $u'\alpha = u\alpha$. The existence of u' is guaranteed by Claim 2.7. Define a word $\omega_0 \in D^*$ by

$$\omega_0 := \bigcirc_{u \in U} (u \rightarrow u')(u' \rightarrow u\alpha).$$

Sort the strong components of D in topological order: C_1, \dots, C_k , i.e. for $i \neq j$, C_i connects to C_j only if $j > i$. For each $1 \leq i \leq k$, define

$$S_i := \{v \in C_i \setminus (U \cup U') : v\alpha \in C_i\},$$

where $U' := \{u' : u \in U\}$, and consider the transformation $\beta_i : C_i \rightarrow C_i$ defined by

$$x\beta_i = \begin{cases} x\alpha & \text{if } x \in S_i \\ x & \text{otherwise.} \end{cases}$$

If $|C_i| \leq 2$ or $C_i \cong P_3$, then $\text{cycl}(\beta_i) = 0$ and β_i can be computed with $|C_i| - \text{fix}(\beta_i)$ arcs. Otherwise, C_i is a complete undirected graph. If $\beta_i \in \text{Sing}(C_i)$, then by Theorem 1.1, there is a word $\omega_i \in C_i^* \subseteq D^*$ of length $|C_i| + \text{cycl}(\beta_i) - \text{fix}(\beta_i)$ expressing β_i . Suppose now that β_i is a non-identity permutation of C_i . By Claim 2.6, α does not permute C_i and there exists $h_i \in C_i \setminus (C_i\alpha)$. Note that $h_i \in C_i \setminus S_i$. Define $\hat{\beta}_i \in \text{Sing}(C_i)$ by

$$x\hat{\beta}_i = \begin{cases} x\alpha & \text{if } x \in S_i \\ a_i & \text{if } x = h_i \\ x & \text{otherwise,} \end{cases}$$

where a_i is any vertex in S_i . Then $\alpha|_{S_i} = \hat{\beta}_i|_{S_i}$. Again by Theorem 1.1, there is a word $\omega_i \in C_i^* \subseteq D^*$ of length $|C_i| + \text{cycl}(\hat{\beta}_i) - \text{fix}(\hat{\beta}) = |C_i| + \text{cycl}(\beta_i) - \text{fix}(\beta_i)$ expressing $\hat{\beta}_i$.

The following word maps all the vertices in $[n] \setminus (U \cup U' \cup C_i)$ that have image in C_i :

$$\omega'_i = \bigcirc \{ (a \rightarrow a\alpha) : a \in [n] \setminus (U \cup U' \cup C_i), a\alpha \in C_i \}.$$

Finally, let

$$\omega := \omega_0 \omega_k \omega'_k \dots \omega_1 \omega'_1 \in D^*.$$

It is easy to check that ω indeed expresses α . Since $\sum_{i=1}^k \text{fix}(\beta_i) = \text{fix}(\alpha) + \sum_{i=1}^k |C_i \setminus S_i|$ and $\sum_{i=1}^k \ell(\omega'_i) = \sum_{i=1}^k |C_i \setminus (U \cup U' \cup S_i)|$, we have

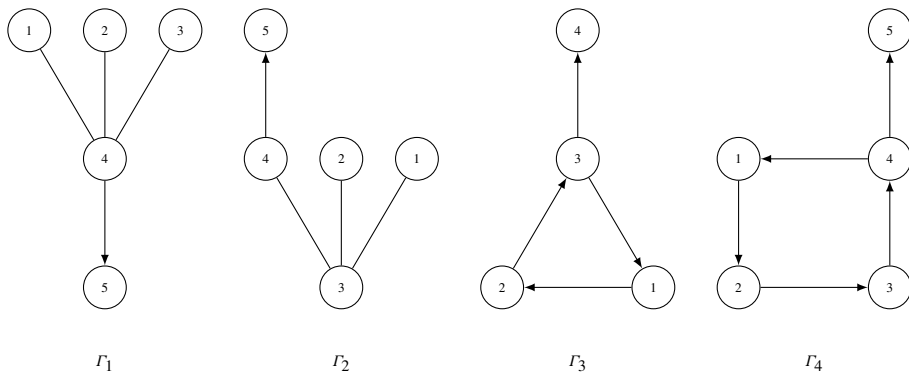
$$\ell(\omega) = 2|U| + \sum_{i=1}^k (\ell(\omega_i) + \ell(\omega'_i)) = n + \sum_{i=1}^k \text{cycl}(\beta_i) - \text{fix}(\alpha) = g(\alpha).$$

□

2.2 Digraphs satisfying condition (C2)

The characterisation of connected digraphs satisfying condition (C2) is based on the classification of connected digraphs D such that $\text{cycl}(\alpha) = 0$, for all $\alpha \in \langle D \rangle$.

For $k \geq 3$, let Θ_k be the directed cycle of length k . Consider the digraphs Γ_1 , Γ_2 , Γ_3 and Γ_4 as illustrated below:



Lemma 2.8 *Let D be a connected digraph on $[n]$. The following are equivalent:*

- (i) *For all $\alpha \in \langle D \rangle$, $\text{cycl}(\alpha) = 0$.*
- (ii) *D has no subdigraph isomorphic to Γ_1 , Γ_2 , Γ_3 , Γ_4 , or Θ_k , for all $k \geq 5$.*

Proof In order to prove that (i) implies (ii), we show that if Γ is equal to Γ_i or Θ_k , for $i \in [4]$, $k \geq 5$, then there exists $\alpha \in \langle \Gamma \rangle$ such that $\text{cycl}(\alpha) \neq 0$.

- If $\Gamma = \Gamma_1$, take

$$\alpha := (3 \rightarrow 4)(4 \rightarrow 5)(1 \rightarrow 4)(4 \rightarrow 3)(2 \rightarrow 4)(4 \rightarrow 1)(3 \rightarrow 4)(4 \rightarrow 2) \\ = 21555.$$

- If $\Gamma = \Gamma_2$, take

$$\alpha := (3 \rightarrow 4)(4 \rightarrow 5)(1 \rightarrow 3)(3 \rightarrow 4)(2 \rightarrow 3)(3 \rightarrow 1)(4 \rightarrow 3)(3 \rightarrow 2) \\ = 21555.$$

- If $\Gamma = \Gamma_3$, take

$$\alpha := (3 \rightarrow 4)(2 \rightarrow 3)(1 \rightarrow 2)(3 \rightarrow 1) = 2144.$$

- If $\Gamma = \Gamma_4$, take

$$\alpha = (3 \rightarrow 4)(4 \rightarrow 5)(2 \rightarrow 3)(3 \rightarrow 4)(1 \rightarrow 2)(4 \rightarrow 1) = 21555.$$

- Assume $\Gamma = \Theta_k$ for $k \geq 5$. Consider the following transformation of $[k]$:

$$(u \Rightarrow v) := (u \rightarrow u_1) \dots (u_{d-1} \rightarrow v),$$

where $u, u_1, \dots, u_{d-1}, v$ is the unique path from u to v on the cycle Θ_k . Take

$$\alpha := (1 \Rightarrow k-3)(k \Rightarrow k-4)(k-1 \Rightarrow 1)(k-2 \Rightarrow k) \\ (k-3 \Rightarrow k-1)(k-4 \Rightarrow k-2).$$

Then, $\alpha = (k-1)(k-1) \dots (k-1) k 1 (k-2)$, where $(k-1)$ appears $k-3$ times, has the cyclic component $(k-2, k)$.

Conversely, assume that D satisfies (ii). If $n \leq 3$, it is clear that $\text{cycl}(\alpha) = 0$, for all $\alpha \in \langle D \rangle$, so suppose $n \geq 4$. We first obtain some key properties about the strong components of \bar{D} .

Claim 2.9 *Any strong component of \bar{D} is an undirected path, an undirected cycle of length 3 or 4, or a claw $K_{3,1}$ (i.e. a bipartite undirected graph on $[4] = [3] \cup \{4\}$). Moreover, if a strong component of D is not an undirected path, then it is terminal.*

Proof Let C be a strong component of \bar{D} . Clearly, C is undirected and, by (ii), it cannot contain a cycle of length at least 5. If C has a cycle of length 3 or 4, then the whole of C must be that cycle and C is terminal (otherwise, it would contain Γ_3 or Γ_4 , respectively). If C has no cycle of length 3 and 4, then C is a tree. It can only be a path or $K_{3,1}$, for otherwise it would contain Γ_1 or Γ_2 ; clearly, $K_{3,1}$ may only appear as a terminal component. \square

Suppose there is $\alpha \in \langle D \rangle$ that has a cyclic orbit (so $\text{cycl}(\alpha) \neq 0$). This cyclic orbit must be contained in a strong component C of \bar{D} , and Claim 2.9 implies that $C \cong \Gamma$, where $\Gamma \in \{K_{3,1}, \bar{\Theta}_s, P_r : s \in \{3, 4\}, r \in \mathbb{N}\}$. If $\Gamma = K_{3,1}$ or $\Gamma = \bar{\Theta}_s$, then C is a terminal component, so α acts on C as some transformation $\beta \in \langle \Gamma \rangle$; however, it is easy to check that no transformation in $\langle \Gamma \rangle$ has a cyclic orbit. If $\Gamma = P_r$, for some r , then α acts on C as a partial transformation β of P_r . Since $\langle P_r \rangle = O_r$, β has no cyclic orbit. \square

We introduce a new property of a connected digraph D :

- ($\star\star$) For every strong component C of D , $|C| \leq 2$ if C is non-terminal, and $|C| \leq 3$ if C is terminal.

Lemma 2.10 *Let D be a closed connected digraph on $[n]$ satisfying property (\star). The following are equivalent:*

- (i) D satisfies property ($\star\star$).
- (ii) D has no subdigraph isomorphic to $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$, or Θ_k , for some $k \geq 5$.

Proof If (i) holds, it is easy to check that D does not contain any subdigraphs isomorphic to $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$, or Θ_k for some $k \geq 5$.

Conversely, suppose that (ii) holds. Let C be a strong component of D . If C is non-terminal, Lemma 2.2 implies that C is complete; hence, $|C| \leq 2$ as otherwise D would contain Γ_4 as a subdigraph. If C is terminal, Lemma 2.2 implies that C is complete or P_3 ; hence, $|C| \leq 3$ as otherwise D would contain Γ_3 as a subdigraph. \square

Theorem 2.11 *Let D be a connected digraph on $[n]$. The following are equivalent:*

- (i) For all $\alpha \in \langle D \rangle$, $\ell(D, \alpha) = n - \text{fix}(\alpha)$.
- (ii) D is closed satisfying properties (\star) and ($\star\star$).

Proof Clearly, D satisfies (i) if and only if it satisfies condition (C1) and $\text{cycl}(\alpha) = 0$, for all $\alpha \in \langle D \rangle$. By Theorem 2.5 and Lemmas 2.8 and 2.10, D satisfies (i) if and only if D satisfies (ii). \square

2.3 Digraphs satisfying condition (C3)

The following result characterises digraphs satisfying condition (C3).

Theorem 2.12 *Let D be a connected digraph on $[n]$. The following are equivalent:*

- (i) For every $\alpha \in \langle D \rangle$, $\ell(D, \alpha) = n - \text{rk}(\alpha)$.
- (ii) $\langle D \rangle$ is a band, i.e. every $\alpha \in \langle D \rangle$ is idempotent.
- (iii) Either $n = 2$ and $D \cong K_2$, or there exists a bipartition $V_1 \cup V_2$ of $[n]$ such that $(i_1, i_2) \in E(D)$ only if $i_1 \in V_1, i_2 \in V_2$.

Proof Clearly (i) implies (ii): if $\ell(D, \alpha) = n - \text{rk}(\alpha)$, then $\text{rk}(\alpha) = \text{fix}(\alpha)$ by inequality (1), so α is idempotent.

Now we prove that (ii) implies (iii). If there exist $u, v, w \in [n]$ pairwise distinct such that $(u, v), (v, w) \in E(D)$, then $\alpha = (v \rightarrow w)(u \rightarrow v)$ is not an idempotent.

Therefore, for $n \geq 3$, if every $\alpha \in \langle D \rangle$ is idempotent, then a vertex in D either has in-degree zero or out-degree zero: this corresponds to the bipartition of $[n]$ into V_1 and V_2 .

We finally prove that (iii) implies (i). Let $n \geq 3$ and suppose that there exists a bipartition $V_1 \cup V_2$ of $[n]$ such that $(i_1, i_2) \in E(D)$ only if $i_1 \in V_1, i_2 \in V_2$. Then for any $\alpha \in \langle D \rangle$, all elements of V_2 are fixed by α and $i_1\alpha \in \{i_1\} \cup N^+(i_1)$ for any $i_1 \in V_1$. In particular, any non-fixed point of α is mapped to a fixed point, so $r := \text{rk}(\alpha) = \text{fix}(\alpha)$. Let $J := \{v_1, \dots, v_{n-r}\} \subseteq V_1$ be the set of non-fixed points of α ; therefore

$$\alpha = (v_1 \rightarrow v_1\alpha) \dots (v_{n-r} \rightarrow v_{n-r}\alpha),$$

where each one of the $n - r$ arcs above belongs to $\langle D \rangle$. The result follows by inequality (1). \square

3 Arc-generated semigroups with long words

Fix $n \geq 2$. In this section, we consider digraphs D that maximise $\ell(D, r)$ and $\ell(D)$. For $r \in [n - 1]$, define

$$\begin{aligned}\ell_{\max}(n, r) &:= \max \{ \ell(D, r) : V(D) = [n] \}, \\ \ell_{\max}(n) &:= \max \{ \ell(D) : V(D) = [n] \}.\end{aligned}$$

The first few values of $\ell_{\max}(n, r)$, calculated with the GAP package *Semigroups* [7], are given in Table 1. By Lemma 2.1, $\ell_{\max}(n, 1) = n - 1$ for all $n \geq 2$; henceforth, we shall always assume that $n \geq 3$ and $r \in [n - 1] \setminus \{1\}$.

In the following sections, we restrict the class of digraphs that we consider in the definition of $\ell_{\max}(n, r)$ and $\ell_{\max}(n)$ to two important cases: acyclic digraphs and strong tournaments.

3.1 Acyclic digraphs

For any $n \geq 3$, let Acyclic_n be the set of all acyclic digraphs on $[n]$, and, for any $r \in [n - 1]$, define

Table 1 First values of $\ell_{\max}(n, r)$

n	r				
	1	2	3	4	5
2	1				
3	2	6			
4	3	11	13		
5	4	18	24	33	
6	5	26	42	51	66

$$\begin{aligned}\ell_{\max}^{\text{Acyclic}}(n, r) &:= \max \{ \ell(A, r) : A \in \text{Acyclic}_n \}, \\ \ell_{\max}^{\text{Acyclic}}(n) &:= \max \{ \ell(A) : A \in \text{Acyclic}_n \}.\end{aligned}$$

Without loss of generality, we assume that any acyclic digraph A on $[n]$ is topologically sorted, i.e. $(u, v) \in E(A)$ only if $v > u$.

In this section, we establish the following theorem.

Theorem 3.1 *For any $n \geq 3$ and $r \in [n - 1] \setminus \{1\}$,*

$$\begin{aligned}\ell_{\max}^{\text{Acyclic}}(n, r) &= \frac{(n - r)(n + r - 3)}{2} + 1, \\ \ell_{\max}^{\text{Acyclic}}(n) &= \ell_{\max}^{\text{Acyclic}}(n, 2) = \frac{1}{2}(n^2 - 3n + 4).\end{aligned}$$

First of all, we settle the case $r = n - 1$, for which we have a finer result.

Lemma 3.2 *Let $n \geq 3$ and $A \in \text{Acyclic}_n$. Then, $\ell(A, n - 1)$ is equal to the length of a longest path in A . Therefore,*

$$\ell_{\max}^{\text{Acyclic}}(n, n - 1) = n - 1.$$

Proof Let v_1, \dots, v_{l+1} be a longest path in A . Then $\alpha \in \langle A \rangle$ defined by

$$v\alpha := \begin{cases} v_{i+1} & \text{if } v = v_i, i \in [l], \\ v & \text{otherwise,} \end{cases}$$

has rank $n - 1$ and requires at least l arcs, since it moves l vertices.

Conversely, let $\alpha \in A$ be a transformation of rank $n - 1$, and consider a word expressing α in A^* :

$$\alpha = (u_1 \rightarrow v_1)(u_2 \rightarrow v_2) \dots (u_s \rightarrow v_s).$$

Since α has rank $n - 1$, we must have $v_2 = u_1$ and by induction $v_i = u_{i-1}$ for $2 \leq i \leq s$. As A is acyclic, $u_s, u_{s-1}, \dots, u_1, v_1$ forms a path in A , so $s \leq l$. \square

The following lemma shows that the formula of Theorem 3.1 is an upper bound for $\ell_{\max}^{\text{Acyclic}}(n, r)$.

Lemma 3.3 *For any $n \geq 3$ and $r \in [n - 1] \setminus \{1\}$,*

$$\ell_{\max}^{\text{Acyclic}}(n, r) \leq \frac{(n - r)(n + r - 3)}{2} + 1.$$

Proof Let A be an acyclic digraph on $[n]$, let $\alpha \in \langle A \rangle$ be a transformation of rank $r \geq 2$, and let $L \subset V(A)$ be the set of terminal vertices of A . For any $u, v \in [n]$, denote the length of a longest path from u to v in A as $\psi_A(u, v)$.

Claim 3.4 $\ell(A, \alpha) \leq \sum_{v \in [n]} \psi_A(v, v\alpha)$.

Proof Let $\omega = (a_1 \rightarrow b_1) \dots (a_l \rightarrow b_l)$ be a shortest word expressing α in A^* , with $l = \ell(A, \alpha)$. Say that the arc $(a_i \rightarrow b_i)$, $i \geq 2$, carries $v \in [n]$ if $v(a_1 \rightarrow b_1) \dots (a_{i-1} \rightarrow b_{i-1}) = a_i$ (assume that $a_1 \rightarrow b_1$ only carries a_1). Every arc $(a_i \rightarrow b_i)$ carries at least one vertex, for otherwise we could remove that arc from the word ω and obtain a shorter word still expressing α . Let $v \in [n]$, and denote $v_0 = v$ and $v_i = v(a_1 \rightarrow b_1) \dots (a_i \rightarrow b_i)$ (and hence $v_l = v\alpha$). Let us remove the repetitions in this sequence: let $j_0 = 0$ and for $i \geq 1$, $j_i = \min\{j : v_j \neq v_{j_{i-1}}\}$. Then the sequence $v = v_{j_0}, v_{j_1}, \dots, v_{j_{l(v)}} = v\alpha$ forms a path in A of length $l(v)$, and hence $l(v) \leq \psi(v, v\alpha)$. For each $v \in [n]$, there are $l(v)$ arcs in ω carrying v , so the length of ω satisfies

$$l \leq \sum_{v=1}^n l(v) \leq \sum_{v \in [n]} \psi_A(v, v\alpha).$$

□

Claim 3.5 If $|L| \geq 2$, then $\sum_{v \in [n]} \psi_A(v, v\alpha) \leq \frac{(n-r)(n+r-3)}{2}$.

Proof As $|L| \geq 2$, and A is topologically sorted, we have $\{n, n-1\} \subseteq L$, and any $\alpha \in \langle A \rangle$ fixes both $n-1$ and n , i.e. $\psi_A(v, v\alpha) = 0$ for $v \in \{n-1, n\}$. For any $v \in [n-2]$, we have

$$\psi_A(v, v\alpha) \leq \min\{n-1, v\alpha\} - v.$$

Hence

$$\begin{aligned} \sum_{v \in [n]} \psi_A(v, v\alpha) &= \sum_{v \in [n-2]} \psi_A(v, v\alpha) \\ &\leq \sum_{v \in [n-2]} (\min\{n-1, v\alpha\} - v) \\ &= \sum_{w \in [n-2]\alpha} \left(\min\{n-1, w\} |w\alpha^{-1}| \right) - T_{n-2}, \end{aligned}$$

where $T_k = \frac{k(k+1)}{2}$. The summation is maximised when $|n\alpha^{-1}| = n-r$ and $|w\alpha^{-1}| = 1$ for $n-r+1 \leq w \leq n-2$, thus yielding

$$\begin{aligned} \sum_{v \in [n]} \psi_A(v, v\alpha) &\leq (n-1)(n-r) + (T_{n-2} - T_{n-r}) - T_{n-2} \\ &= \frac{(n-r)(n+r-3)}{2}. \end{aligned}$$

□

Claim 3.6 If $|L| = 1$, then $\ell(A, \alpha) \leq \frac{(n-r)(n+r-3)}{2} + 1$.

Proof As A is topologically sorted, $L = \{n\}$. We use the notation from the proof of Claim 3.4. We then have $l(n) = 0$. We have three cases:

- Case 1** $(n-1)$ is fixed by α . Then, $l(n-1) = 0$ and $l(v) \leq \min\{n-1, v\alpha\} - v$ for all $v \in [n-2]$. By the same reasoning as in Claim 3.5, we obtain $\ell(A, \alpha) \leq \frac{(n-r)(n+r-3)}{2}$.
- Case 2** $(n-1)\alpha = n$ and $v\alpha \leq n-1$ for every $v \in [n-2]$. Then again $l(v) \leq \min\{n-1, v\alpha\} - v$, for all $v \in [n-2]$, and $\ell(A, \alpha) \leq \frac{(n-r)(n+r-3)}{2}$.
- Case 3** n has at least two pre-images under α . Let $\omega = (a_1 \rightarrow b_1) \dots (a_l \rightarrow b_l)$ be a shortest word expressing α in A^* , and denote $\alpha_0 = \text{id}$ and $\epsilon_i = (a_i \rightarrow b_i)$, $\alpha_i = \epsilon_1 \dots \epsilon_i$ for $i \in [l]$. We partition $n\alpha^{-1}$ into two parts S and T :

$$S = \{v \in n\alpha^{-1} : v_{l(v)-1} = n-1\}, \quad T = n\alpha^{-1} \setminus S.$$

For all $v \in S$, if the arc carrying v to $n-1$ is ϵ_j , then $(n-1)\alpha_{j-1}^{-1} \subseteq S$ (v can only collapse with other pre-images of α). Then the arc $(n-1 \rightarrow n)$ occurs only once in the word ω (if it occurs multiple times, then remove all but the last occurrence of that arc to obtain a shorter word expressing α). If we do not count that arc, we have $l'(v) \leq n-1-v$ arcs carrying v if $v \in S$, $l(v) \leq n-1-v$ arcs carrying v if $v \in T$, and $l(v) \leq v\alpha - v$ if $v\alpha \neq n$. Again, we obtain $\ell(A, \alpha) \leq \frac{(n-r)(n+r-3)}{2} + 1$. \square

Lemma 3.3 follows by the previous claims. \square

The following lemma completes the proof of Theorem 3.1.

Lemma 3.7 *For any $n \geq 3$ and $r \in [n-1] \setminus \{1\}$, there exists an acyclic digraph Q_n on $[n]$ and a transformation $\beta_r \in \langle Q_n \rangle$ of rank r such that*

$$\ell(Q_n, \beta_r) \geq \frac{(n-r)(n+r-3)}{2} + 1.$$

Proof Let Q_n be the acyclic digraph on $[n]$ with edge set

$$E(Q_n) := \{(u, u+1) : u \in [n-1]\} \cup \{(n-2, n)\}.$$

For any $r \in [n-1] \setminus \{1\}$, define $\beta_r \in \langle Q_n \rangle$ by

$$v\beta_r := \begin{cases} n-r+v & \text{if } v \in [r-2], \\ n-1 & \text{if } v \in [n-1] \setminus [r-2], \ n-v \equiv 0 \pmod{2}, \\ n & \text{if } v \in [n-1] \setminus [r-2], \ n-v \equiv 1 \pmod{2}, \\ n & \text{if } v = n. \end{cases}$$

Let β_r be expressed as a word in Q_n^* of minimum length as

$$\beta_r = (a_1 \rightarrow b_1) \dots (a_l \rightarrow b_l),$$

where $l = \ell(Q_n, \beta_r)$. Denote $\alpha_0 := \text{id}$, $\epsilon_i := (a_i \rightarrow b_i)$, and $\alpha_i := \epsilon_1 \dots \epsilon_i$, for $i \in [l]$. Say that ϵ_i carries $u \in [n]$ if $u\alpha_{i-1} = a_i$ and hence $u\alpha_i \neq u\alpha_{i-1}$.

Claim 3.8 For each $i \in [l]$, the arc ϵ_i carries exactly one vertex.

Proof First, $(a_1, b_1) \in E(Q_n)$ and $a_1\beta_r = b_1\beta_r$ imply that $a_1 = n - 1$ and $b_1 = n$. Suppose that there is an arc ϵ_j , $j \in [l]$, that carries two vertices $u < v$; take j to be minimal index with this property. We remark that $v \leq n - 2$ and $u\alpha_{j-1} = v\alpha_{j-1}$ imply $u\beta_r = v\beta_r$. Then $w := u + 1$ satisfies $w\beta_r \neq u\beta_r$, so w is not carried by ϵ_j . If $w\alpha_{j-1} \leq n - 2$, then $u\alpha_{j-1} < w\alpha_{j-1} < v\alpha_{j-1}$ since $u < w < v$ and the graph induced by $[n - 2]$ in Q_n is the directed path \vec{P}_{n-2} ; this contradicts that $u\alpha_{j-1} = v\alpha_{j-1}$. Hence $w\alpha_{j-1} \geq n - 1$ and $v\alpha_{j-1} \geq n - 1$. If $v\alpha_{j-1} = n$ or $v\beta_r = n - 1$, then ϵ_j does not carry v . Thus, $v\alpha_{j-1} = n - 1$ and $v\beta_r = n$. Then, in order to carry v to $n - 1$, we have $\epsilon_s = (n - 2 \rightarrow n - 1)$ for at least one $s \in [l]$, and $\epsilon_j = (n - 1 \rightarrow n)$. For $s \in [j - 1]$, replace all occurrences $\epsilon_s = (n - 2 \rightarrow n - 1)$ with $\epsilon'_s := (n - 2 \rightarrow n)$ and delete ϵ_j : this yields a word in Q_n^* of length $l' < l$ expressing β_r , which is a contradiction. \square

For all $i \in [l]$, denote $\delta(i) := \sum_{v \in [n]} d_{Q_n}(v\alpha_i, v\beta_r)$. We then have $\delta(l) = 0$, and by the claim, $\delta(i) \geq \delta(i - 1) - 1$ for all $i \in [l]$. Thus $l \geq \delta(0)$, where

$$\begin{aligned} \delta(0) &= \sum_{v \in [n]} d_{Q_n}(v, v\beta_r) \\ &= \sum_{v=1}^{r-2} (n - r) + \sum_{v=r-1}^{n-2} (n - 1 - v) + 1 \\ &= \frac{(n - r)(n + r - 3)}{2} + 1. \end{aligned}$$

\square

3.2 Strong tournaments

Let $n \geq 3$. Recall that if T is a strong tournament on $[n]$, then $\{a \rightarrow b : (a, b) \in E(T)\}$ is a minimal generating set of Sing_n . Let Tour_n denote the set of all strong tournaments on $[n]$. For $r \in [n - 1]$, define

$$\begin{aligned} \ell_{\max}^{\text{Tour}}(n, r) &:= \max\{\ell(T, r) : T \in \text{Tour}_n\}, \\ \ell_{\max}^{\text{Tour}}(n) &:= \max\{\ell(T) : T \in \text{Tour}_n\}. \end{aligned}$$

Define analogously $\ell_{\min}^{\text{Tour}}(n, r)$ and $\ell_{\min}^{\text{Tour}}(n)$. The first few values of $\ell_{\min}^{\text{Tour}}(n, r)$ and $\ell_{\max}^{\text{Tour}}(n, r)$, calculated with the GAP package *Semigroups* [7] using data from [6], are given in Table 2. The calculation of these values has been the inspiration for the results of this section and the conjectures of the next one.

Lemma 3.9 Let $n \geq 3$ and $T \in \text{Tour}_n$.

1. For any partition P of $[n]$ into r parts, there exists an idempotent $\alpha \in \text{Sing}_n$ with $\ker(\alpha) = P$ such that $\ell(T, \alpha) = n - r$.

Table 2 First values of $(\ell_{\min}^{\text{Tour}}(n, r), \ell_{\max}^{\text{Tour}}(n, r))$

n	r				
	2	3	4	5	6
3	(6, 6)				
4	(8, 8)	(11, 11)			
5	(6, 11)	(8, 14)	(10, 17)		
6	(8, 13)	(10, 18)	(11, 21)	(13, 24)	
7	(8, 16)	(10, 22)	(11, 26)	(13, 29)	(15, 32)

2. For any r -subset S of $[n]$, there exists an idempotent $\alpha \in \text{Sing}_n$ with $\text{Im}(\alpha) = S$ such that $\ell(T, \alpha) = n - r$.

Proof 1. Let $P = \{P_1, \dots, P_r\}$. For all $1 \leq i \leq r$, the digraph $T[P_i]$ induced by P_i is a tournament, so it is connected and there exists a vertex v_i reachable by any other vertex in P_i : let α map the whole of P_i to v_i . Then α , when restricted to P_i , is a constant map, which can be computed using $|P_i| - 1$ arcs. Summing for i from 1 to r , we obtain that $\ell(T, \alpha) = n - r$.

2. Without loss of generality, let $S = [r] \subseteq [n]$. For every $v \in [n]$, define

$$s(v) := \min\{s \in S : d_T(s', v) \geq d_T(s, v), \forall s' \in S\}.$$

In particular, if $v \in S$, then $s(v) = v$. Moreover, if $v = v_0, v_1, \dots, v_d = s(v)$ is a shortest path from v to $s(v)$, with $d = d_T(v, s(v))$, then $s(v_i) = s(v)$ for all $0 \leq i \leq d$. For each $v \in [n]$, fix a shortest path P_v from v to $s(v)$, and consider the digraph D on $[n]$ with edges

$$E(D) := \{(a, b) : (a, b) \in E(P_v) \text{ for some } v \in [n]\}.$$

Then, D is acyclic and the set of vertices with out-degree zero in D is exactly S . Let sort $[n]$ so that D has reverse topological order: $(a, b) \in E(D)$ only if $a > b$. Note that S is fixed by this sorting. Let α be given by $v\alpha := s(v)$; hence, with the above sorting

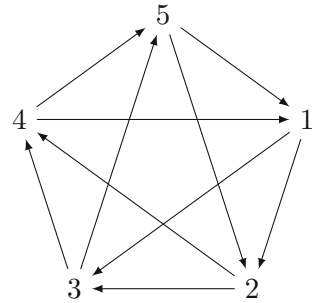
$$\alpha = \bigcirc_{v=n}^{r+1}(v \rightarrow v_1).$$

□

Lemma 3.10 Let $n \geq 3$, $T \in \text{Tour}_n$, and $\alpha := (u \rightarrow v) \in \text{Sing}_n$, for $(u, v) \notin E(T)$. Then

$$\ell(T, \alpha) = 4d_T(u, v) - 2.$$

Proof Let $u = v_0, v_1, \dots, v_d = v$ be a shortest path from u to v in T , where $d := d_T(u, v)$. As $(u, v) \notin E(T)$ and T is a tournament, we must have $(v, u) \in E(T)$.

Fig. 2 Circulant tournament κ_5 

By the minimality of the path, for any $j + 1 < i$, we have $(v_j, v_i) \notin E(T)$, so $(v_i, v_j) \in E(T)$. Then, the following expresses α with arcs in T^* :

$$\begin{aligned} (v_0 \rightarrow v_d) = & (v_d \rightarrow v_0)(v_{d-1} \rightarrow v_d)(v_{d-2} \rightarrow v_{d-1}) \cdots (v_1 \rightarrow v_2)(v_0 \rightarrow v_1) \\ & ((v_2 \rightarrow v_0)(v_1 \rightarrow v_2))((v_3 \rightarrow v_1)(v_2 \rightarrow v_3)) \\ & \cdots ((v_d \rightarrow v_{d-2})(v_{d-1} \rightarrow v_d)) \\ & (v_{d-2} \rightarrow v_{d-1}) \cdots (v_0 \rightarrow v_1). \end{aligned}$$

So $\ell(T, \alpha) \leq 4d - 2$. For the lower bound, we note that any word in T^* expressing $(u \rightarrow v)$ must begin with $(v \rightarrow u)$. Then, u has to follow a walk in T towards v ; say this walk has length $l \geq d$. All the vertices on the walk must be moved away (as otherwise they would collapse with u) and have to come back to their original position (since α fixes them all); as the shortest cycle in a tournament has length 3, this process adds at least $3(l - 1)$ symbols to the word. Altogether, this yields a word of length at least

$$1 + l + 3(l - 1) = 4l - 2 \geq 4d - 2.$$

□

Let $n = 2m + 1 \geq 3$ be odd, and let κ_n be the *circulant tournament* on $[n]$ with edges $E(\kappa_n) := \{(i, (i + j) \bmod n) : i \in [n], j \in [m]\}$. Figure 2 illustrates κ_5 . In the following theorem, we use κ_n to provide upper and lower bounds for $\ell_{\min}^{\text{Tour}}(n, r)$ and $\ell_{\max}^{\text{Tour}}(n, r)$ when n is odd.

Theorem 3.11 *For any n odd, we have*

$$\begin{aligned} n + r - 2 &\leq \ell_{\min}^{\text{Tour}}(n, r) \leq n + 8r, \\ (\hat{r} + 1)(n - \hat{r}) - 1 &\leq \ell_{\max}^{\text{Tour}}(n, r) \leq 6rn + n - 10r. \end{aligned}$$

where $\hat{r} = \min\{r - 1, \lfloor n/2 \rfloor\}$.

Proof Let $T \in \text{Tour}_n$ and $2 \leq r \leq n - 1$. We introduce the following notation:

$$[n]_r := \{\mathbf{u} := (u_1, \dots, u_r) : u_i \neq u_j, \forall i, j\},$$

$$\Delta(T, r) := \max \left\{ \sum_{i=1}^r d_T(u_i, v_i) : \mathbf{u}, \mathbf{v} \in [n]_r \right\}.$$

The result follows by the next claims.

Claim 3.12 $r'(\text{diam}(T) - r' + 1) + r - r' \leq \Delta(T, r) \leq r \text{diam}(T)$, where $r' = \min\{r, \lfloor (\text{diam}(T) + 1)/2 \rfloor\}$.

Proof The upper bound is clear. For the lower bound, let $u, v \in [n]$ be such that $d_T(u, v) = \text{diam}(T)$, and let $u = v_0, v_1, \dots, v_d = v$ be a shortest path from u to v , where $d = \text{diam}(T)$. Then, $d_T(v_i, v_j) = j - i$, for all $0 \leq i \leq j \leq D$. If $1 \leq r \leq \lfloor (d + 1)/2 \rfloor$, consider $\mathbf{u}' = (v_0, \dots, v_{r-1})$ and $\mathbf{v}' = (v_{d-r+1}, \dots, v_d)$, so we obtain $\Delta(T, r) \geq r(d - r + 1)$. If $r \geq \lfloor (d + 1)/2 \rfloor$, simply add vertices u'_j and v'_j such that $(u'_j, v'_j) \notin T$. \square

Claim 3.13 $\min\{\Delta(T, r) : T \in \text{Tour}(n)\} = \Delta(\kappa_n, r) = 2r$.

Proof Let $\mathbf{u} = (u_1, \dots, u_n)$ form a Hamiltonian cycle, and choose $\mathbf{v} = (u_n, u_1, \dots, u_{n-1})$. Then $d_T(u_i, v_i) \geq 2$ for all i . Conversely, since $\text{diam}(\kappa_n) = 2$, we have $\Delta(\kappa_n, r) = 2r$. \square

Claim 3.14 $n - r + \Delta(T, r - 1) \leq \ell(T, r) \leq n + 6r \text{diam}(T) - 4r$.

Proof For the lower bound, consider $\alpha \in \text{Sing}_n$ as follows. Let $\mathbf{u} = (u_1, \dots, u_{r-1})$ and $\mathbf{v} = (v_1, \dots, v_{r-1})$ achieve $\Delta(T, r - 1)$, and let $v \notin \{v_1, \dots, v_{r-1}\}$; define

$$x\alpha = \begin{cases} v_i & \text{if } x = u_i, \\ v & \text{otherwise.} \end{cases}$$

Let $\omega = e_1 \dots e_l$ (where $e_i = (a_i \rightarrow b_i)$) be a shortest word expressing α , where $l := \ell(T, \alpha)$. Recall that an arc e_i carries a vertex c if $ce_1 \dots e_{i-1} = a_i$. By the minimality of ω , every arc carries at least one vertex. Moreover, if c and d are carried by e_i , then $c\alpha = d\alpha$; therefore, we can label every arc e_i of ω by an element $c(e_i) \in \text{Im}(\alpha)$ if e_i carries vertices eventually mapping to $c(e_i)$. Denote the number of arcs labelled c as $l(c)$, we then have $l = \sum_{c \in \text{Im}(\alpha)} l(c)$. For any $u \in V$, there are at least $d_T(u, u\alpha)$ arcs carrying u . Therefore,

$$l = \sum_{c \in \text{Im}(\alpha)} l(c) \geq \sum_{i=1}^{r-1} d_T(u_i, v_i) + \sum_{a \notin \mathbf{u}} d_T(a, v) \geq \Delta(T, r - 1) + n - r.$$

For the upper bound, we can express any $\alpha \in \text{Sing}_n$ of rank r in the following fashion. By Lemma 3.9, there exists $\beta \in \text{Sing}_n$ with the same kernel as α such that

$\ell(T, \beta) = n - r$. Suppose that $\text{Im}(\alpha) = \{v_1, \dots, v_r\}$ and $\text{Im}(\beta) = \{u_1, \dots, u_r\}$, where $u_i \beta^{-1} = v_i \alpha^{-1}$, for $i \in [r]$. Let $h \in [n] \setminus \text{Im}(\beta)$. Define a transformation γ of $[n]$ by

$$x\gamma = \begin{cases} v_i & \text{if } x = u_i, \\ v_1 & \text{if } x = h, \\ x & \text{otherwise.} \end{cases}$$

Then $\alpha = \beta\gamma$, where $\gamma \in \text{Sing}_n$, and by Theorem 1.1

$$\ell(K_n, \gamma) = n - \text{fix}(\gamma) + \text{cycl}(\gamma) \leq r + \frac{r}{2} = \frac{3r}{2}.$$

By Lemma 3.10, each arc associated with K_n may be expressed in at most $4\text{diam}(T) - 2$ arcs associated with T ; therefore,

$$\ell(T, \gamma) \leq \frac{3r}{2}(4\text{diam}(T) - 2) = 6r\text{diam}(T) - 3r.$$

Thus,

$$\ell(T, \alpha) \leq \ell(T, \beta) + \ell(T, \gamma) \leq n + 6r\text{diam}(T) - 4r.$$

□
□

4 Conjectures and open problems

We finish the paper by proposing few conjectures and open problems.

Let π_n be the tournament on $[n]$ with edges $E(\pi_n) := \{(i, (i + 1) \bmod n) : i \in [n]\} \cup \{(i, j) : j + 1 < i\}$. Figure 3 illustrates π_5 .

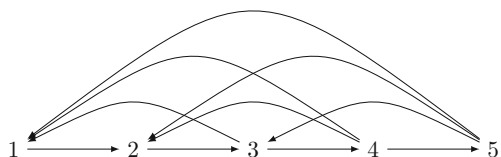
Conjecture 4.1 *For every $n \geq 3$, $r \in [n - 1]$, and $T \in \text{Tour}_n$, we have*

$$\ell(T, r) \leq \ell(\pi_n, r) = \ell_{\max}^{\text{Tour}}(n, r),$$

with equality if and only if $T \cong \pi_n$. Furthermore,

$$\ell(\pi_n) = \ell_{\max}^{\text{Tour}}(n) = \frac{n^2 + 3n - 6}{2},$$

Fig. 3 π_5



which is achieved for $\alpha := n(n-1) \dots 2n$.

Tournament π_n has appeared in the literature before: it is shown in [8] that π_n has the minimum number of strong subtournaments among all strong tournaments on $[n]$. On the other hand, it was shown in [1] that, for n odd, the circulant tournament κ_n has the maximal number of strong subtournaments among all strong tournaments on $[n]$.

Conjecture 4.2 For every $n \geq 3$ odd, $r \in [n-1]$, and $T \in \text{Tour}_n$, we have

$$\ell_{\min}^{\text{Tour}}(n, r) = \ell(\kappa_n, r).$$

Furthermore,

$$\ell_{\min}^{\text{Tour}}(n, 2) = n + 1 \text{ and } \ell_{\min}^{\text{Tour}}(n, r) = n + r,$$

for all $3 \leq r \leq \frac{n+1}{2}$.

Conjecture 4.3 There exists $c > 0$ such that for every simple digraph D on $[n]$, $\ell(D) = O(n^c)$.

The referee of this paper noted that the automorphism groups of K_n and $\langle K_n \rangle = \text{Sing}_n$ are both isomorphic to Sym_n and proposed the following problems.

Problem 1 Investigate connections between the automorphism groups of D and $\langle D \rangle$. Is it possible to classify all digraphs D such that the automorphism group of D and of $\langle D \rangle$ are isomorphic?

Problem 2 Generalise the ideas of this paper to oriented matroids. Is there a natural way to associate (not necessarily idempotent) transformations to each signed circuit of an oriented matroid?

In a forthcoming paper, we investigate the relationship between the graph theoretic properties of D and the semigroup properties of $\langle D \rangle$.

Acknowledgments The second and third authors were supported by the EPSRC grant EP/K033956/1. We kindly thank the insightful comments and suggestions for open problems of the anonymous referee of this paper.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Beineke, L.W., Harary, F.: The maximum number of strongly connected subtournaments. Can. Math. Bull. **8**, 491–498 (1965)
2. Howie, J.M.: The subsemigroup generated by the idempotents of a full transformation semigroup. J. Lond. Math. Soc. **41**, 707–716 (1966)

3. Howie, J.M.: Idempotent generators in finite full transformation semigroups. *Proc. R. Soc. Edinb.* **81A**, 317–323 (1978)
4. Howie, J.M.: Products of idempotents in finite full transformation semigroups. *Proc. R. Soc. Edinb.* **86A**, 243–254 (1980)
5. Iwahori, N.: A length formula in a semigroup of mappings. *J. Fac. Sci. Univ. Tokyo Sect. 1A Math.* **24**, 255–260 (1977)
6. McKay, B.: Catalogue of Directed Graphs. September 2015. Retrieved from <https://cs.anu.edu.au/people/Brendan.McKay/data/digraphs.html>
7. Mitchell, J.D., et al.: Semigroups—GAP Package, Version 3.0. September 2015
8. Moon, J.W.: On subtournaments of a tournament. *Can. Math. Bull.* **9**, 297–301 (1966)
9. Solomon, A.: Catalan monoids, monoids of local endomorphisms, and their presentations. *Semigroup Forum* **53**, 351–368 (1996)
10. You, T., Yang, X.: A classification of the maximal idempotent-generated subsemigroups of finite singular groups. *Semigroup Forum* **64**, 236–242 (2002)
11. Yang, X., Yang, H.: Maximal regular subsemibands of Sing_n . *Semigroup Forum* **72**, 75–93 (2006)
12. Yang, X., Yang, H.: Isomorphisms of transformation semigroups associated with simple digraphs. *Asian Eur. J. Math.* **2**(4), 727–737 (2009)